家蚕神经肽基因的筛查及成熟肽的预测

甘　玲，梁九波，刘喜龙，何宁佳*
（西南大学蚕学与系统生物学研究所，农业部蚕桑学重点开放实验室，重庆 400716）

摘要：神经肽(neuropeptide)是家蚕 *Bombyx mori* 体内重要的调控物质。为了充分理解神经肽对家蚕发育的调控，扩充家蚕神经肽的数量，本研究利用 BLAST 的 tblastn 程序结合 OpenOffice 软件的查找程序，基于其他昆虫和无脊椎动物神经肽的同源性和保守的结构特点，在家蚕基因、理论蛋白质数据库及 NCBI 中进行全面而系统的基因筛查；并利用各种在线工具对所筛查到基因和理论蛋白质的结构和成熟肽进行预测分析。结果共获得 allatostatin-A (AST-A), allatostatin-B (AST-B), allatostatin-C (AST-C), allatropin (AT), ecdysis-triggering hormone (ETH), crustacean cardioactive peptide (CCAP)和 FMRFamide 等 31 个神经肽基因家族，包括 37 个神经肽基因亚家族，共计 44 个神经肽基因；预测出 193 个成熟的神经肽，其中 73 个根据家族同源性预测在 C 末端发生酰胺化，而 6 个被预测在 N 末端发生了环化，9 个被预测酪氨酸发生了硫酸化。大部分成熟神经肽都具有明显的家族结构特征，但 proctolin, CCAP 及 CAPA-PK 成熟肽结构上与其他昆虫相比有所扩展。结果提示，家蚕神经和内分泌细胞产生了几乎在所有昆虫中具有的神经肽前体；进化过程中大部分成熟神经肽的氨基酸序列在种间产生了差异，但家族特征性基序高度保守。本研究为神经肽功能研究以及神经肽对家蚕发育尤其是蛹期发育调控的研究奠定了基础。

关键词：家蚕；神经肽；基因筛查；成熟肽；酰胺化；环化；硫酸化；信息分析
中图分类号：Q966　**文献标识码**：A　**文章编号**：0454-6296(2010)01-0009-11

Neuropeptide gene screening and mature peptide prediction in the silkworm, *Bombyx mori*

GAN Ling, LIANG Jiu-Bo, LIU Xi-Long, HE Ning-Jia* (Key Sericulture Laboratory of the Ministry of Agriculture, Institute of Sericulture and Systems Biology, Southwest University, Chongqing 400716, China)

Abstract: Neuropeptides play an important role of regulation for the silkworm, *Bombyx mori*. To fully comprehend the regulation of neuropeptides in the growth of *B. mori* and to get more neuropeptides in *B. mori*, the online program tblastn of BLAST and the search program of OpenOffice software were used to screen genes encoding putative neuropeptide precursors in the databases of silkworm genome and theoretical proteins based on the homology and the structure of the neuropeptides conserved in other insects and invertebrates. Several online programs were used to analyze the structures of the genes and the theoretical proteins to predict the mature peptides. As results, 31 neuropeptide gene families with 37 gene subfamilies including allatostatin-A (AST-A), allatostatin-B (AST-B), allatostatin-C (AST-C), allatropin (AT), bombyxin, ecdysis-triggering hormone (ETH), crustacean cardioactive peptide (CCAP), and FMRFamide were identified. A total of 44 neuropeptide genes and 193 mature peptides were predicted. Among them, 73 neuropeptides are amidated at the C-terminal, six neuropeptides are cyclized at the N- terminal, and nine neuropeptides are sulfated. Most of the mature peptides share the similarity because they belong to the same family. But compared with that of other insects, mature peptides of proctolin, CCAP and CAPA-PK are more advanced in the structure. The results suggest that the neuropeptide precursors, which are possessed by almost all the insects, are produced by nerve and endocrine cells of *B. mori*, and in the course of evolution, differences in the amino acid sequences for most mature peptides formed among different insect species, but the motifs for the same family are highly conserved. This study provided useful information for future research on functions of neuropeptides, which will give us novel insights into the regulation of developmental process, especially at

---

基金项目：国家高技术研究发展计划("863"计划)项目(2006AA10A118)
作者简介：甘玲，女，1978 年生，四川资中人，讲师，博士研究生，研究方向为生物化学与分子生物学，E-mail: gl9089@tom.com
*通讯作者 Corresponding author, E-mail: hejia@swu.edu.cn
收稿日期 Received: 2009-09-04；接受日期 Accepted: 2009-12-18

the pupal stage, in the silkworm.

Key words: *Bombyx mori*; neuropeptide; gene screening; mature peptide; amidation; cyclization; sulfation; information analysis

蚕丝业起源于中国，历史悠久，为中华民族文化经济和社会发展作出了极为重要的贡献。蚕丝业将在解决"三农"问题，特别是增加农民收入、调整农村产业结构上发挥越来越重要的作用。家蚕是由古代野桑蚕驯化而成的绢丝昆虫。由于人们有目的地长期选择和培育，使蚕体内的丝腺非常发达，产丝能力强，丝质也好。家蚕是完全变态的昆虫，在它的一生中，要经历卵、幼虫、蛹、成虫（蛾）4 个在形态和功能上完全不同的发育阶段。研究表明昆虫神经肽（neuropeptide）是昆虫体内重要的调控因素，是昆虫体内最大量和最多样化的信号分子，在调控昆虫的生长发育、采食、繁殖等方面起着关键性的作用（Kastin, 2006）。尤其是生长抑制肽（growth-blocking peptide, GBP）对鳞翅目昆虫生长发育调控的研究，引起了科研工作者的关注。这些肽分子通常来自于前体蛋白，通过在特定序列处的酶切和翻译后修饰加工成为成熟的具有生物活性的神经肽（Kastin, 2006）。家蚕中分泌神经肽的器官主要是脑、咽下神经节及其他神经节内的神经分泌细胞。为了充分理解神经肽对家蚕蛹期的发育调控，对编码神经肽基因的筛查是必须的。随着不同昆虫基因组测序计划的完成和公开，加速了神经肽基因的发现（Coates *et al.*, 2000; Hauser *et al.*, 2008），为昆虫的发育、行为学和进化关系等都提供了重要的信息。在家蚕中，前人的研究已证实有几种神经肽在家蚕内分泌机制或生长发育的调控上起重要的作用，如促前胸腺激素（PTTH）、羽化激素（EH）等，但仅限于对单个神经肽基因的研究。由于生物体是一个复杂的网络，任何一个刺激都会牵动网络的许多环节，只对少数几个神经肽进行研究显然是不够的。随着家蚕基因组计划的实施和完成，一个庞大的基因数据库已经建成。本研究旨在利用已知的昆虫或无脊椎动物神经肽基因信息和家蚕的基因组数据库、蛋白质数据库，通过保守基序筛查家蚕中可能存在的神经肽基因，并进行生物信息学分析，为家蚕神经肽分子及生物学功能研究以及神经肽在发育调控中的重要作用及机制研究奠定基础。

1 材料和方法

1.1 数据库搜索

根据在其他昆虫如黑腹果蝇 *Drosophila melanogaster*、冈比亚按蚊 *Anopheles gambiae*、意大利蜜蜂 *Apis mellifera*、烟草天蛾 *Manduca sexta* 以及赤拟谷盗 *Tribolium castaneum* 等多种昆虫及无脊椎动物中已经研究报道的神经肽保守基序，利用 BLAST 的 tblastn 程序，结合 OpenOffice 软件的查找程序，在家蚕基因组数据库、预测蛋白质数据库中进行全面而系统的筛查。特别是利用在线程序 tblastn（http://silkworm.swu.edu.cn/silksoft/blast2-simple.html）（Altschul *et al.*, 1997）进行神经肽基因的搜索（blast score 范围值为 32 ~ 168；本地 blast 搜索，E 值设置为 6E-004），并利用 ExPASy-PROSITE 软件（http://www.expasy.org/prosite/）（Sigrist *et al.*, 2002）、ClustalX1.81 软件、NCBI 的 psi-blast 程序（http://blast.ncbi.nlm.nih.gov/Blast.cgi）（Altschul *et al.*, 1997）、MEME server（http://meme.sdsc.edu/meme/）等对不同家族亚家族理论神经肽前体全序列进行比对分析。

1.2 成熟肽结构的预测

利用 ExPASy 的 Translate tool（http://www.expasy.ch/tools/dna.html）（Gasteiger *et al.*, 2003）对筛查得到的神经肽基因进行核酸的翻译。利用在线软件 Signal P3.0 的 Neural Networks 和 Hidden Markov Models 算法程序（http://www.cbs.dtu.dk/services/SignalP/）（Bendtsen *et al.*, 2004）进行神经肽前体信号肽的预测。前体肽的切割位点利用预测工具 NeuroPred（http://neuroproteomics.scs.uiuc.edu/neuropred.html）（Southey *et al.*, 2006）的 Known Motif 和 Insect Models 方法并结合 Veenstra（2000）提供的信息以及与已知节肢动物神经肽前体加工程序或与神经肽基因家族已知成熟神经肽的同源性进行预测；使用在线软件 MAFFT 版本 6（http://align.bmr.kyushu-u.ak.jp/mafft/online/server/）（Katoh and Toh, 2008）对推测的氨基酸进行序列比对；使用在线程序 Sulfinator（http://www.expasy.org/tools/sulfinator）（Monigatti *et al.*, 2002）结合与已知节肢动物神经肽的同源性进行酪氨酸残基硫酸化状态的预测；N 末端谷氨酰胺/谷氨酸的环化及 C 末端甘氨酸残基的酰胺化等根据与已知肽亚型的同源性进行预测。

2　结果与分析

2.1　家蚕神经肽基因的特点

表 1 中列出了家蚕神经肽基因筛查的结果。此外本研究对几个最近被证实和测序但不知其前体的十足类肽也进行了搜索，即 HI/LGSI/LYRamide 和-Ryamides（Stemmler *et al*.,2007），但在家蚕中没有找到同源的编码基因。表 1 中未列出的家蚕素（bombyxin）是结构与胰岛素（insulin）/松弛素（relaxin）相关的肽，在家蚕基因组中家蚕素基因有 32 个拷贝，根据它们序列的相似性被划分为 6 个家族（Yoshida *et al*., 1998）。

表 1　筛查到的家蚕神经肽基因

Table 1　Neuropeptide genes in *Bombyx mori*

神经肽家族 Neuropeptide family	亚家族 Subfamily	基因编号 Number in database	基因组的区域 Genomic region	信号肽 Signal peptide	已知的功能 Known functions
AKH		BGIBMGA008722-TA	Bm_nscaf2912	No	参与幼虫血液碳水化合物动态平衡的控制
		BGIBMGA008672-TA	Bm_nscaf2912	Yes	
		gi:195182593	Bm_nscaf684	Yes	
Apis-ITG-like		gi:193248555	Bm_nscaf2623	Yes	
AST-A		BGIBMGA014377-TA	Bm_scaffold416	Yes	
AST-B		BGIBMGA009464-TA	Bm_nscaf2953	Yes	抑制前胸腺蜕皮甾醇的合成
AST-C		BGIBMGA005369-TA	Bm_nscaf2828	No	
AT		BGIBMGA011850-TA	Bm_nscaf3031	Yes	
BMK5		BGIBMGA008978-TA	Bm_nscaf2930	Yes	
Bursicon	Bursicon α	gi:62529358	Bm_nscaf3015	Yes	参与家蚕翅的扩展
	Bursicon β	BGIBMGA011086-TA	Bm_nscaf3015	Yes	
CAPA		gi:191175143	Bm_nscaf2954	Yes	
		gi:191175145	Bm_nscaf2954	Yes	
CCAP		BGIBMGA009606-TA	Bm_nscaf2962	Yes	
CCHamide		gi:194072591	Bm_scaffold701	Yes	
CHH	CHH	BGIBMGA004168-TA	Bm_nscaf2770	Yes	
Corazonin		BGIBMGA002280-TA	Bm_nscaf2330	Yes	减少纺丝率，抑制化蛹
DH	Calcitonin-like	BGIBMGA011588-TA	Bm_nscaf3027	Yes	
	Corticotropin-releasing factor-like	gi:193248573	Bm_nscaf2888	Yes	
		gi:193248575	Bm_nscaf2888	Yes	
		gi:193248577	Bm_nscaf2888	Yes	
EH		BGIBMGA006291-TA	Bm_nscaf2852	Yes	引发蜕皮或羽化
ETH		BGIBMGA009728-TA	Bm_nscaf2964	Yes	促进蜕皮激素释放
	FMRFamide	BGIBMGA009271-TA	Bm_nscaf2943	Yes	抑制前胸腺蜕皮素生成
	Myosuppressin	BGIBMGA010073-TA	Bm_nscaf2983	Yes	对脏腑肌有抑制作用；抑制前胸腺的活性
FMRFamide-related peptide	NPF	BGIBMGA009803-TA	Bm_nscaf2964	Yes	
		gi:191175137	Bm_nscaf2852	Yes	
		gi:40861544	Bm_nscaf2852	Yes	
	sNPF	BGIBMGA012522-TA	Bm_nscaf3048	No	
	Sulfakinin	BGIBMGA013558-TA	Bm_nscaf3078	Yes	
GBP		BGIBMGA014551-TA	Bm_scaffold802	Yes	
GPA2		gi:194072586	Bm_nscaf2998	Yes	
GPB5		BGIBMGA010342-TA	Bm_nscaf2990	Yes	

续表 1 Table 1 continued

神经肽家族 Neuropeptide family	亚家族 Subfamily	基因编号 Number in database	基因组的区域 Genomic region	信号肽 Signal peptide	已知的功能 Known functions
IMFamide		gi:193248554	Bm_nscaf2964	Yes	
Kinin		BGIBMGA010075-TA	Bm_nscaf2983	Yes	
Neuroparsin		BGIBMGA006409-TA	Bm_nscaf2853	Yes	
NPLP1		BGIBMGA009024-TA	Bm_nscaf2930	Yes	
Orcoknin		BGIBMGA005422-TA	Bm_nscaf2828	No	
PBAN		BGIBMGA001651-TA	Bm_nscaf2176	Yes	胚胎期的滞育
PDH		gI:193248572	Bm_nscaf2859	Yes	
Proctolin		BGIBMGA008345-TA	Bm_nscaf2902	No	
PTTH		BGIBMGA000357-TA	Bm_nscaf1681	No	促进前胸腺合成和释放蜕皮甾体
SIFamide		gi:193248552	Bm_nscaf2964	Yes	
Tachykinin		BGIBMGA010076-TA	Bm_nscaf2983	Yes	

AKH：脂肪动员激素 Adipokinetic hormone；AST：咽侧体抑制素 Allatostatin；AT：促咽侧体素 Allatotropin；CAPA：能量肽 Capability；CCAP：甲壳类刺激心脏肽 Crustacean cardioactive peptide；CHH：甲壳动物高血糖激素 Crustacean hyperglycemic hormone；DH（Cal-like）：降钙素样利尿激素 Calcitonin-like diuretic hormone；DH（Cor-releasing factor-like）：促肾上腺皮质激素样利尿激素 Corticotropin-releasing factor-like diuretic hormone；EH：羽化激素 Eclosion hormone；ETH：蜕皮释放激素 Ecdysis-triggering hormone；GBP：生长抑制肽 Growth-blocking peptide；GPA2：糖蛋白激素 α2 Glycoprotein hormone alpha2；GPB5：糖蛋白激素 β5 Glycoprotein hormone beta5；NPF：神经肽 F Neuropeptide F；NPLP1：神经肽样前体 1 Neuropeptide like precursor 1；PBAN：信息素生物合成激活肽 Pheromone biosynthesis activating neuropeptide；PDH：色素分散激素 Pigment dispersing hormone；PTTH：促前胸腺激素 Prothoracicotropic hormone；sNPF：短神经肽 F Short neuropeptide F. 表 2 同 The same for Table 2. 表中"基因组的区域"中的 Bm 代表 *Bombyx mori* Bm in the column of genomic region represents *Bombyx mori*；表中"信号肽"栏中的"Yes"表示具有信号肽，"No"表示没有预测到信号肽 Yes and No mean with and without signal peptide, respectively. 表中的"基因编号"栏，数据库是指家蚕基因组数据库(http://silkworm.swu.edu.cn/silk db/)和 NCBI，带"gi"的编号来自 NCBI，其余的编号来自家蚕基因组数据库。Database means silkworm databank (http://silkworm.swu.edu.cn/silk db/) and NCBI. The numbers with gi are from NCBI, and the rest from silkworm databank.

预测的神经肽前体序列与在另外的昆虫中报道的神经肽前体相比较表明，家蚕神经和内分泌细胞产生了几乎所有昆虫中具有的前体神经肽，除了加压素样肽(vasopressin-like peptide)和神经肽样前体 2(neuropeptide-like precursor 2，NPLP2)。家蚕神经肽基因的保守性主要体现在以下几个方面：(1)成熟肽在前体中的位置保守。大部分神经肽前体都具有如下的结构特征：信号肽＋成熟短神经肽＋(前体相关肽)，比如 ETH；(2)许多重复的具有典型家族特征的成熟肽序列被碱基切割位点分开，如 FMRFamide 相关肽、AST-B、CAPA 家族等。当然在不同昆虫的同源物中，重复数是不相同的；(3)基因的选择性剪接模式是保守的，在家蚕中如 DH 基因转录后的选择性剪接在另外的昆虫中也存在相同方式(Dai *et al*., 2007)。

2.2 成熟神经肽的预测

肽构成多细胞生物体内最大和最多样化的信号分子。这些分子通常来自于更大的蛋白质，即肽前体蛋白，通过在特定序列的酶切加工后，许多肽经历广泛的翻译后修饰，再形成其成熟的、具有生物学活性的构型。神经肽前体的生物激活被几个不同的机制产生，如在成对的碱性氨基酸处切割成更小的肽；在羧基端酰胺化，天冬酰胺糖基化，酪氨酸残基硫酸化，以及谷氨酰胺环化形成焦谷氨酸(Eipper *et al*., 1986)。这些过程对保证成熟肽的生物学活性和稳定性具有重要的作用。

以 B 型咽侧体抑制素(allatostatin-B，AST-B)为例，阐述神经肽基因的搜索及成熟肽的预测过程(图 1)。利用果蝇的 B 型咽侧体抑制素前体肽(GenBank accession no.：NP-648971)进行搜索，家蚕 BGIBMGA009464-TA 基因(blast score：69；E-value：8e-13)为假定的编码 B 型咽侧体抑制素神经肽前体基因。其编码的理论蛋白 BGIBMGA009464-PA 为假定的 B 型咽侧体抑制素神经肽前体(图 1)。Signal P3.0 的 Neural Networks 和 Hidden Markov Models 算法程序预测表明这个前体肽序列前 19 个氨基酸作为信号肽，信号肽酶在 Ala^{19} 和 Ala^{20} 之间进行酶切处理，剩下的序列包含了 17 个 Lys-Arg 和 Lys-Lys 的二碱基酶切加工位点，其中 12 个位点之前有 1 个 Gly 残基，表明了各自切割产物 C 末端的酰胺化。除了信号肽序列外，预测出 18 个假定的肽，其中有 12 个符合 B 型咽侧体抑制素家族 C 末端保守基序的结构 WX6Wamide(X 表示在两个 Trp 残基之间的 6 个可变氨基酸)，即 AWQDMSSAWamide，AWQDLNSAWamide，GWQDLNS

AWamide，AWQDLNSAWamide，GWQDLNSAWamide，SWQDLNSVWamide，AWQDLNSAWamide，AWQDLNSAWamide，GWNDMSSAWamide，AWQDLNSAWamide，AWSALHGTWamide 和 AWQDMSSAWGKQAPEKWAAFHGSWamide(图 1 和表 1)。剩下的 6 个肽序列，即 AEEPHHDAAPQPYKTSICALRSAVLPH，PDKCSTNSYKTRFPVTGSNNHTLNTDNEVDLTEDD，AWSSLHSGWA，DDEAME，PVKPMFNNGGY，SSIEPDY$_{(SO3H)}$EEIDAVEQLVPYQQVPNEEHIDAPE 为 B 型咽侧体抑制素前体相关肽(图 1 和表 1)。

PROTEIN:BGIBMGA009464-PA

MRWCLFALWVFGVATVVTAAEEPHHDAAPQPYKTSICALRSAVLPHKKPDKCSTNSYKTRFPVTGSNNHTLNTDNEVDLTEDDKRAWSSL
HSGWAKRAWQDMSSAWGKRAWQDLNSAWGKRGWQDLNSAWGKRAWQDLNSAWGKRGWQDLNSAWGKRDDDEAMEKKSWQDLNS
VWGKRAWQDLNSAWGKRAWQDLNSAWGKRGWNDMSSAWGKRAWQDLNSAWGKRAWQDMSSAWGKQAPEKWAAFHGSWGKRSSI
EPDYEEIDAVEQLVPYQQVPNEEHIDAPEKKAWSALHGTWGKRPVKPMFNNGGY

↓

AEEPHHDAAPQPYKTSICALRSAVLPHKKPDKCSTNSYKTRFPVTGSNNHTLNTDNEVDLTEDDKRAWSSLHSGWAKRAWQDMSSAWGK
RAWQDLNSAWGKRGWQDLNSAWGKRAWQDLNSAWGKRGWQDLNSAWGKRDDDEAMEKKSWQDLNSVWGKRAWQDLNSAWGKRA
WQDLNSAWGKRGWNDMSSAWGKRAWQDLNSAWGKRAWQDMSSAWGKQAPEKWAAFHGSWGKRSSIEPDYEEIDAVEQLVPYQQVPN
EEHIDAPEKKAWSALHGTWGKRPVKPMFNNGGY

↓

AEEPHHDAAPQPYKTSICALRSAVLPHKK　PDKCSTNSYKTRFPVTGSNNHTLNTDNEVDLTEDDKR　AWSSLHSGWAKR
AWQDMSSAWGKRAWQDLNSAWGKR GWQDLNSAWGKR AWQDLNSAWGKR GWQDLNSAWGKR DDDEAMEKK SWQDLNSVWGKR
AWQDLNSAWGKR　AWQDLNSAWGKR　GWNDMSSAWGKR　AWQDLNSAWGKR　AWQDMSSAWGKQAPEKWAAFHGSWGKR
SSIEPDYEEIDAVEQLVPYQQVPNEEHIDAPEKK　AWSALHGTWGKR　PVKPMFNNGGY

↓

AEEPHHDAAPQPYKTSICALRSAVLPH;
PDKCSTNSYKTRFPVTGSNNHTLNTDNEVDLTEDD;
AWSSLHSGWA AWQDMSSAWG AWQDLNSAWG GWQDLNSAWG AWQDLNSAWG GWQDLNSAWG　DDDEAME　SWQDLNSVWG
AWQDLNSAWG　AWQDLNSAWG GWNDMSSAWG AWQDLNSAWG AWQDMSSAWGKQAPEKWAAFHGSWG
SSIEPDYEEIDAVEQLVPYQQVPNEEHIDAPE　AWSALHGTWG
PVKPMFNNGGY

↓

AWQDMSSAWamide　AWQDLNSAWamide　GWQDLNSAWamide　AWQDLNSAWamide　GWQDLNSAWamide
SWQDLNSVWamide　AWQDLNSAWamide　AWQDLNSAWamide　GWNDMSSAWamide　AWQDLNSAWamide
AWQDMSSAWGKQAPEKWAAFHGSWamide
SSIEPDYEEIDAVEQLVPYQQVPNEEHIDAPE　AWSALHGTWamide

↓

SSIEPDY$_{(SO3H)}$EEIDAVEQLVPYQQVPNEEHIDAPE

图 1　家蚕一个假定的 B 型咽侧体抑制素(AST-B)前体激活过程

Fig. 1　The precursor activation process of a putative B-type allatostatin (AST-B) in *Bombyx mori*

图中家蚕 BGIBMGA009464-TA 基因编码了一个假定的前体 A putative precursor is encoded by silkworm BGIBMGA009464-TA gene in the figure. 前体假定的信号肽序列用斜体表示 The putative precursor signal peptide sequence is shown in italics. 图 2 同 The same for Fig. 2. allatostatin-B (AST-B) 前体假定的加工流程表示如下：第 1 个箭头代表信号肽酶的切割；第 2 个箭头代表假定的前体转化酶的切割；第 3 个箭头代表通过羧肽酶 E 的酶切加工，假定的肽序列被释放；第 4 个箭头代表肽酰甘氨酸 α-酰胺化单氧化酶对 C 末端甘氨酸残基的酰胺化；第 5 个箭头代表肽序列中酪氨酸残基的硫酸化作用。The processes of B-type allatostatins (AST-B) are shown with arrows: the first arrow stands for the cutting of signal peptidase; the second stands for the cutting of the putative precursor invertase; the third stands for the release of the putative peptide sequence by carboxypeptidase E; the fourth stands for the amidation of peptidyl-glycine α amidation monooxygenase to C-terminal glycine residues; and the fifth stands for the sulfation of the tyrosine residue in the peptide sequence.

其余的神经肽前体用相同的方法和思路进行预测，预测出的成熟肽见表2。软件分析结果显示，5个神经肽基因没有预测到信号肽序列，暗示这些基因5′末端不完整。另外也可能是在5′末端到第一个蛋氨酸密码子之间有一段测序错误，因为典型的神经肽信号序列在这一段富含非极性和疏水性氨基酸残基。序列的错误或不完整可能导致预测的成熟肽数量减少或结构不准确。因此可通过与家族其他成员的同源性进行比较预测。

PROTEIN:BGIBMGA008722-PA

MAEAQLTFTSSWGGKRAAIAGTVSCRNDESLASIYKLIQNEAEKLLLCQKP

图2 家蚕一个假定的AKH前体成熟肽的预测

Fig. 2 The prediction of a putative AKH precursor mature peptide in *Bombyx mori*

尽管信号肽预测表明 BGIBMGA008722-TA 编码前体(图2)没有信号肽存在，但所有已知的AKH前体有如下相似的结构组成特点：信号肽+单拷贝成熟肽+标准的酰胺和二碱基切割信号(-GKR或者-GRR)+AKH前体相关肽(APRP)。编码肽pQLNFSPGWamide的几个前体蛋白通常作为AKH/RPCH家族。所有的AKHs成员在N末端谷氨酸环化形成焦谷氨酸，而C末端酰胺化(Fernlund, 1974)，因此可以推测在Ala4与Gln5之间为信号肽酶切割位点，预测出该基因编码的家蚕典型的AKH成熟肽为pQLTFTSSWGamide，前体相关肽结构为AAIAGTVSCRNDESLASIYKLIQNEAEKLLLCQKP。

2.3 神经肽的结构

2.3.1 神经肽结构的家族保守性：许多成熟神经肽都具有明显的家族结构特征，这些结构基序在强的进化压力下高度保守。因此既可结合同源性比对进行神经肽基因的筛查，也可对成熟肽的预测结果进行验证。预测结果表明，大部分神经肽基因编码前体可切割成具有典型家族特征的家族神经肽和前体相关肽(表2)，其中73个根据家族同源性预测在C末端发生酰胺化，而6个被预测在N末端发生了环化，9个被预测在酪氨酸被硫酸化。由此可见家蚕神经肽前体切割，C末端发生酰胺化是最普遍的转录后修饰方式。其中预测得到3个具有AKH/RPCH家族典型特征的成熟肽结构，即pQLTFTSSWGamide(九肽)、pQLTFTPGWGQamide(十肽)、pQITFSRDWSGamide(十肽)。在家蚕中一个与烟草天蛾AKH完全相同的九肽pQLTFTSSWGamide已经通过化学的方法得到了证实(Ishizaki and Suzuki, 1992)。SIFamide是一个高度保守的昆虫/甲壳类动物神经调节肽，拥有XYRKPPFNGSIFamide的结构，位点1处的X是一个可变的残基，预测得到的家蚕成熟肽SIFamide结构为TYRKPPFNGSIFamide，与该家族成员相比，仅在第一个氨基酸残基有差异。IMFamide与SIFamide前体同源性非常高(39%的相同性)且都位于同一基因组nscaf2964 contig中。IMFamide基因与SIFamide基因具有相反的转录方向。预测得到家蚕的corazonin成熟肽结构为pQTFQYSRGWTNamide，与最早从美洲蜚蠊 *Periplaneta americana* 的心脂肪体分离得到的corazonin结构完全相同(Veenstra, 1989)。昆虫甲壳类高血糖激素/昆虫离子转运肽(Crustacean hyperglycemic hormone, CHH/insect ion transport peptide, ITP)属于肽的一个大家族，在甲壳动物中的生理作用是显著的。根据前体相关肽的存在与否，可被划分为两个亚家族。即CHH亚家族和蜕皮抑制激素(MIH)亚家族。其中CHH亚家族含有一个前体相关肽，而MIH亚家族则没有。本研究对该家族基因成熟肽的预测结果SFFTLECKGVFDAAIFARLDICDDCFNLFREPQLYTLCRAECFTTPYFKGCMESLYLYDEKEQIDQMIDFVamide与烟草天蛾中描述的离子转运肽SFFSLECKGVYDASIFARLDRVCDDCYNLFREPQLYTLCRAKCFTTPYFKGCMESLYLYDEKEQIDQMIDFVamide(Dai *et al.*, 2007)有高度的同源性，并且在前体中有一个前体相关肽，是一个真正的CHH亚家族成员。预测得到的FMRFamide相关肽家族的5个亚家族成员具有保守的C末端保守的结构特征和酰胺化的残基(表2)。在烟草天蛾中，拥有pQVRFRQCYFNPISCF序列的肽已被表明具有强的咽侧体活性抑制作用(Kramer *et al.*, 1991)。此后，对另外几种昆虫的与pQVRFRQCYFNPISCF相关的肽特征也得到了描述，如果蝇的pQVRYRQCYFNPISCF(Williamson *et al.*, 2001)。这些结果已形成了一个结构相关肽家族，共同的称为C型咽侧体抑制素(AST-C)。对其家蚕C型咽侧体抑制素预测结果产生的pQVRFRQCYFNPISCF与烟草天蛾完全相同。NFDEIDRSGFGFN肽是最早从刺额小龙虾的神经系统分离且描述其特征。表明在后肠具有刺激肌肉运动的作用。根据它的来源及生物学活性，这个肽被命名为orcoknin(Stangier *et al.*, 1992)。orcoknin的另一个亚型紧接着在大量的十足类及几种昆虫中被

证实（Pascual *et al.*，2004），因此把 N 末端为 NFDEIDR 保守基序的肽归为 orcoknin 肽家族。预测结果表明家蚕含有两个 orcoknin 家族的成熟肽，即 NFDEIDRSSLNTFV 和 NFDEIDRSSMPFPYAI。此外，其他的神经肽基因也通过预测得到了具有保守家族特征的成熟肽。综上所述，尽管在昆虫的进化过程中大部分神经肽的氨基酸序列在种间产生了差异，但家族特征性基序高度保守。

表 2　预测的家蚕成熟神经肽

Table 2　Predicted mature neuropeptides in *Bombyx mori*

神经肽家族（亚家族） Family (subfamily)	预测的成熟肽的结构 Predicted mature peptide structure	家族保守基序 Conservative motif
AKH	pQLTFTSSWGamide, pQLTFTPGWGQamide, pQITFSRDWSGamide	pQX$_6$Wamide
AKH-PRP	AAIAGTVSCRNDESLASIYKLIQNEAEKLLLCQKP, SEATDYRNDGCSSEDSVYTIYKLIKNEAEKFLACQS, HFVVSMQPHHSICSSGPRTQLGSR	
Apis-ITG-like	WGGLFNRFSSDMLANLGYGRSPYRHYPYGQVEPDEAYEALENNRISNVIDEPAHCYSSPCVTNGDC CGGLLCLETDDGGRCLSAFAGRKLGEICNRENQCDAGLICEEAAPGEMHICRPPSTGRKQY NEDCTTSSECEITRGLCCIMQ, HRQKSRKSCGYFKEPLVCIGPVAIDQIREYVEHTTGE, IGAYRLH	
AST- A	SPQYDFGLamide, LPVYNFGLamide, PYSFGLamide, ARMYSFGLamide, ARSYSFGLamide, QRDMHRFSFGLamide, LSSKFNFGLamide	-Y/FXFGLamide
AST-A-PRP	SEDDTSENYIDA, AEQVAEHGAEQVAEHAAPLE, AYSYVSEY, FAEEPAED,	
AST-B	AWQDMSSAWamide, AWQDLNSAWamide, GWQDLNSAWamide AWQDLNSAWamide, GWQDLNSAWamide, SWQDLNSVWamide AWQDLNSAWamide, AWQDLNSAWamide, GWNDMSSAWamide AWQDLNSAWamide, AWSALHGTWamide, AWQDMSSAWGKQAPEKWAAFHGSWamide	-WX$_6$Wamide
AST-B-PRP	AEEPHHDAAPQPYKTSICALRSAVLPH, PDKCSTNSYKTRFPVTGSNNHTLNTDNEVDLTEDD, AW SSLHSGWA, DDDEAME, PVKPMFNNGGY, SSIEPDY$_{(SO3H)}$EEIDAVEQLVPYQQVPNEEHIDAPE	
AST-C	pQVRFRQCYFNPISCF	
AST-C-PRP	MNRVARSWPQATEPRGWAMRNVDGRLVRPWRAD *	
AT	GFKNVEMMTARGFamide, GTPTFKSPSVGIARDFamide, GTFKPNSNILIARGYamide	-ARXYamide
AT-PRP	APDGRLA, KQQRPT, DRSHTRAE, DVDHQRSSQRPS, TPEFENEGIWEADIDLPCVALAIYTRLA, SGNDDYDEEEEIRVT, SILVESKLDFVRDWLSKPQTGPFEKSVSLCRYDIGYVHHPRMKFTPKIGLM VAHDFGamide, TNSQNEPVYGVNNYWDMLENAPEREGQEANDDKTLERKINDRTKVSFRQMFLA NYLLKNDNQNGHDIFFCHLVINYV	
BMK5	GSGTKVGGAGAATKVVTKSGSamide, VSAVPTPSNNKDGSTISELPENWDQTKDDNRSLFLNKSDKNDLEPYPLALSEEGNQDGYDQTVDQR FDSPQSNGELDNLIMRPELYGEPPAMEGLASAFDLQ, NLKPEDQAALSPIDLMTQHEAQ, LIL, GSGTK VGGAAASAKTATKNSGGN, NFRPISE, DSGLSAADVRALLNLWEAQE, FMVA, KQEYANQFA ADRYYGRVNPDEEQPEVDENGDLWYNEPVVIGPHDRDYPHHSYFSEQNRMALARGYPDLYQVG PNELAQRYEEA, QYANKM, SDNMMHQNNYRPRDDLYTLAELLRSAPRVQEQDIPVY	
Bursicon α	FPVTGHEVQLPPGTKFFCQECQMTAVIHVLKHRGCKPKAIPSFACIGKCTSYVQVSGSKIWQMERTC NCCQESGEREATVVLFCPDAQNEE, FRKVSTKAPLQCMCRPCGSIEESSIIPQEVAGYSEEGPLYNHFRKSL	
Bursicon β	EENCETVASEVHVTKEEYDEMGRLLRSCSGEVSVNKCEGMCNSQVHPSISSPTGFQKECFCCREKF LRERLVTLTHCYDPDGIRFEDEENALMEVRLREPDECECYKCGDFSR	
CAPA-pvk	KLRPDGVLNLYPFPRVamide, QLYAFPRVamide	-PRVamide
CAPA-pk	GGPPSDRNEPHDDLLGLHLDDPGMWFGPRLamide, GGPPSDRNEPHDDLLGLHLDDPGMWFGPRLDVVNQNEDamide	-PRLamide
CAPA-PRP	ASYRTWQIPINDVY$_{(SO3H)}$LDY$_{(SO3H)}$EPVE, SEREQIDQIAHEERMK, SKLL, SLKNGDDDV VNQNEDamide	
CCAP	PFCNAFTAVTVTRYSQSFSKNYLSIRRLRALTKPIKCFKHVLFYQGCamide	PFCNAFTGCamide

续表 2 Table 2 continued

神经肽家族(亚家族) Family (subfamily)	预测的成熟肽的结构 Predicted mature peptide structure	家族保守基序 Conservative motif
CCAP-PRP	KQKEFYNSQ, ATIPRNFDPRSNEEMVTMP, SQTAPGMPNQDLMRQRQYVDEDTLGTMLDSESAIDELSRQILSEAKLWEAIQEASAEIA	
CCHamide	GCSAFGHSCFGGHamide	CX_3GHXCXGXHamide
CCHamide-PRP	SQNTNDEY($_{SH03}$)FLETI, SGEPAPMDMANQDMMVRHQLGQEETPPHPGYPHSSYNVLQPGD DIIPIRDGGVYDHDAAARDVMKYKLRNIFKHWMDNY	
CHH	SFFTLECKGVFDAAIFARLDICDDCFNLFREPQLYTLCRAECFTTPYFKGCMESLYLYDEKEQID QMIDFVamide	
CHH-PRP	LPSSAPHHVE	
Corazonin	pQTFQYSRGWTNamide	pQTFQYSRGWTNamide
Corazonin-PRP	YKGERNHELFDVFQ, DGH, DELRDEVLERILTPCQLDKLKYVLEGKPLNDRLFVPCDYIEE EVNQP	
DH(Cal-like)	AFDLGLGRGYSGALQAKHLMGLAAANFAGGPamide	
DH(Cal-like)-PRP	NAQ, YPSINDYYRDDGQYDPDEIIDMLGRLGNLIQMERKMQNE	
DH(Cor-releasing factor-like)	MPSLSIDMPMSVLRQKLSLENERKLQSLRAMANRNFLNDIamide, SFSVNPAVDTLQRGAYNHFLEKVVQSNRDYLNRIamide, KMPSLSINNPMEVLRQRLLLEVARKQMREANQRQAVANRLFLQNVamide	
DH(Cor-releasing factor-like)-PRP	IDLMQMDPSLADDESLGFAMQSLSGRYAAAPWLYLLADVSHDPQNGSDRVK, GFHWAPSAKAAKFY, IDLMQMDPSLADDESLGFAMQSLSGRYAAAPWLYLLADVSHDPQVRSEMQARS, DSWKILTEN, IDLMQ MDPSLADDESLGFAMQSLSGRYAAAPWLYLLADVSHDPQRMAEFSQSSGRARP, GAWGEPASYLYNN	
EH	MFGPWFEGSLCAESCIKARGKDIPECESFASISPFLNKL, SPAIASSYDAMEICIENCAQC	
ETH	SNEAFDEDVMGYVIKSNKNIPRMamide	-PRXamide
PETH	SFIKPNNVPRVamide	-PRXamide
ETH-PRP	NYDSGNHFDIPKVYSLPFEFYGDNEKSLNNDDAEEYYAKKMGSM	
FMRFamide	SAIDRSMIRFamide, SASFVRFamide, DPSFIRFamide, HRARNHFIRLamide	-RXamide
FMRFamid-PRP	SPDLEA, STLPVVPPAQPSFLQRYSASQPAALTADDLMTFLRAYEEDYSSPVS, SVDEENSGYQAETNTYPQ, SVVDPCNDCA, DNELSESNDEDRY$_{(SO3)}$EVESERT	
GBP	GVNGFFNDLRRGISQVEEDLSDRLIFRDDDNDQYNYNNAVNRPVYPTESVSSGSVQRGVVEFVTQ PTIVPTPTSAGKTTTEKEGRENFVGGCATGFTADGRCKPTF	
GPA2	HPVTSLGQCCNIMEAEDVPVKVLCLDGERNLIFKSAVSCACYHCQKE, AESWRKPGCHRIGHTRNISIPDCVEFKITTNACRGYCESWSLPSIMLGF	
GPB5	CWDDVKIVSCWGYCLSYEVCSSEDTSCEWLPPDSSLLGGLILKEELEEELE, FFCVDASMSVRCKL, HSHKVMQTDLNS	
IMFamide	NYKNAPMNGIMFamide	
IMFamide-PRP	GPTDYDSRSKTFTALCEIATEACQAWFPTQEN	
Kinin	NFSPWGamide, VRFSPWGamide, KVKFSAWGamide	-FX1X2WGamide
Kinin-PRP	FPLKMMALSAKFDPRSFKEAMPFKTFVESLPKVFKPGQPYYDVNIKKDamide, SPPIIGPIWTPAPENVKDSTLDTIILIRNSPDKDEAIKTV; DKPSFEHMWTW; SSSVREPSMP, QYINTGGPGADTILDPLDSSLQYSLPYANYFNVPDAQMDPDTSDGQYGIVHDGGVQ, TADDKEQVRD, WLPNLADIDKTMYIKNNEVATPFLVGFSGNSDENDTALLDQVD; SGQMVYKPGSKSSKLIFSATVPELEKIVSNYLPSGERLNLAGLHYIPSVD	
Myosuppressin	pQDVVHSFLRFamide	-HVFLRFamide
Myosuppressin-PRP	APAQLCAAAAENDPRATRFCQALNTFIELYAEAAGEQVPEYQALIRDYPQLLDTGM	
Neuroparsin	WDCVCNPRECEVLEPSGCPGLGIVVWDSCRCCKVCARTLGENCGGFSGTCEPGLKCYEGSCTQIT	

续表 2 Table 2 continued

神经肽家族(亚家族) Family (subfamily)	预测的成熟肽的结构 Predicted mature peptide structure	家族保守基序 Conservative motif
NPF	PERFDTAEQISNYLKELQEYYSVHGRGRYamide, REEGPNNVAEALRILQLLDNYYTQAARPRFamide, DVDAAGDRVDPELLDRAVRLLWLEKLDRIYSYHTRPRFamide	-RXRXamide
NPF-PRP	FGYK; QYPRP; QMHIADASVIFRESPFFEHSLNEDGLL; SDTYTNWAKDVEKPDLPTWLTYA; REEGPNNVAEALRILQLLDNYYTQAARPRID; SDTYTNWAKDVEKPDLPTWLTY A	
NPLP1	NIQALARDGYRMamide, PNFIDAFTKNRWQPITKESPKFYYFRSLKIPLHSSamide, HIGSLAREYFAAPEVDSAVNSDDADV AVPPHSHPTGRILHRPLSNDLPNTNPSLPAPPISP, KNQ, LPTEPENQWPTFP, NIATLAKNGYLRNSGANSY, SISILAKNGMLPTYRSPYVGTDKQEHEDESQE, NLASIARLRSYSAM, GQYNTQNN	
Orcoknin	NFDEIDRSSLNTFV, NFDEIDRSSMPFPYAI	NFDEIDR
Orcoknin-PRP	DPHQEQDVSGNDHTSGERYDAAGHKTRNLDPLGGGNLVRDVREARHSGYLPYQMF, YDYISPYamide, EPWPLVPAEYGGYYGDGFP, FYHLSSFD, YRPDYPMDEIDLSHFPVGS, SQDSY$_{(SO3H)}$ PLASNL	
PBAN	TDMKDESDRGAHSERGALWFGPRLamide, IIFTPKLGRSVAKPQTHESLEFIPRLamide, LSEDMPATPADQEMYQPDPEEMESRTRYFSPRLamide, TMSFSPRLamide	-FXPRL/I/Vamide
PBAN-PRP	ELSYGMFNVLGTFAVRP, SMKPSTEDNRQTFLRLLEAADALKFYYDQLPYERQADEPETKVT	
PDH	NADLINSLLALPKDMNDAamide	
PDH-PRP	VANSEAKIKLNRKVSESSYGSDEQYIRQIHSLVNAYREDNSMLGENFIIETKALIDTKYRTW	
Proctolin	QGSDDETGCIRLLVCKITPFIVRMQEAVFGNNGRNDKTKSKGNEQQRNGVASVMYRYLPT	-RYLPT
Proctolin-PRP	HRDCSIE, QDEINEHSDICE, MTWILNKNSRKLQEMISAAKNESLPDNIEHLI *	
PTTH	PDVGGFMVEDQRTHKSHNYMM, ARNDVLGDKENVRPNPYYTEPFDPDTSPEELSALIVDYANMIRNDVILLDNSVETRT, GNIQVENQAIPDPPCTCKYKKEIEDLGENSVPRFIETRNCNKTQQPTCRPPYICKESLYSITIL, ETKSQESLEIPNELKYRWVAESHPVSVACLCTRDYQLRYNNN	
SIFamide	TYRKPPFNGSIFamide	XYRKPPFNGSIFamide
SIFamide-PRP	NNVENDSSGRAIAALCEITTETCQAWYQALESQ	
sNPF	SVRSPSRRLRFamide, SDPDMPPQAPIDEMDELLSLREVRTPVRLRFamide, SDERAVPHIFPQEEQDRAVRAPSMRLRFamide	-RXRFamide
sNPF-PRP	QALSNYDASPATFESRNNWDALGGLYALLAQHDALGGHALA, SDNNMFLMPYESALPKEVK ASGSVEDDRQQE	
Sulfakinin	GDDTFDDY$_{(SO3H)}$ GHLRFamide	-Y$_{(SO3H)}$ GHM/LRFamide.
Sulfakinin-PRP	SDD, ANL, IQPDDDEDFRPHPLY$_{(SO3H)}$ RDYGLIRSRVI	
Tachykinin	IPQGFLGMRamide, KPQFFVGVKamide, APLGFTGVRamide, AANMHQFYGVRamide, PYDLSIRGKFIGVRamide, GQMGFFGMRamide	-FXGXRamide
Tachykinin-PRP	WIDGMNDPRY, QEMI, HEDDSSEQYY, EDMSSEY$_{(SO3H)}$ QY$_{(SO3H)}$ Y$_{(SO3H)}$ PYEAL, DGSLIGQIEYTSAEHINDGQYPILNDILNEYLQKLERQETNSDTNETEEQRITNEVE, SVDN, DVKNSNGKEIKFLLSRPFP, NFY$_{(SO3H)}$ DY$_{(SO3H)}$ LENPDGY$_{(SO3H)}$ F	

PRP：前体相关肽 Precursor-related peptide；CAPA-PVK：能量肽 pvk Capability-periviscerokinin ；CAPA-PK：能量肽 pk Capability-pyrokinin；* 表示由于信号肽缺失，预测受到影响的成熟肽结构；X 代表可变的氨基酸残基，其后的数字代表不同的变量，其下标的数字代表间隔的氨基酸残基数。氨基酸字母右下标的-(SO3H)-代表酪氨酸残基硫酸化状态(SO3H)；N 末端小写字母 p 代表谷氨酰胺/谷氨酸的环化；C 末端-amide 代表甘氨酸残基的酰胺化。Asterisk indicates the affected prediction of the mature peptide structure due to the missing signal peptide; X stands for the variable amino acid residues, and the subsequent figures represent the different variables. The footnote figures represent the number of the interval amino acid residue. The right footnote of amino acid refers to the state of tyrosine sulfate residue. The N-terminal letter p represents the cyclization of glutamine/glutamate. C-terminal-amide represents the amidation of glycine residues.

2.3.2 神经肽结构的扩展性：预测结果表明家蚕部分神经肽结构存在扩展现象。Proctolin 是一个五肽，即 RYLPT，最早来自于在美洲蜚蠊 *Periplaneta americana*（Starratt and Brown，1975）的描述，在大量的昆虫和十足甲壳类上已经证实了它的结构和功能。proctolin 主要作为神经调质起作用（Isaac *et al.*，2004）。使用黑腹果蝇 proctolin 的前体肽（GenBank accession no：CAD30643）并结合 openoffice 软件的查找程序在家蚕理论蛋白质数据库中进行搜索，BGIBMGA008345-TA 为假定的编码含有 proctolin 的前体基因。但对其前体预测的成熟肽结构为 QGSDDETGCIRLLVCKITPFIVRMQEAVFGNNGRNDKTKSKGNEQQRNGVASVMYRYLPT。与另外同源的成熟亚型比较表明 proctolin N 端存在扩展现象（图 3：A），这种扩展现象在桃蚜 *Myzus persicae* 中也有报道（Christie *et al.*，2008）。此外，含有 PFCNAFTGCamide 结构的肽最早从普通滨蟹 *Carcinus maenas* 的神经系统中分离得到证实，表明具有作用于心脏的作用，因此被命名为作用于心脏的肽或 CCAP（Stangier *et al.*，1987）。利用烟草天蛾的 CCAP 前体肽（GenBank accession no：AAL39064）进行搜索，BGIBMGA009606-TA 被证实为假定的编码 CCAP 的神经肽前体基因（blast score：149；E-value：5e-37）。对其前体预测结果产生的该家族成熟肽结构 PFCNAFTAVTVTRYSQSFSKNYLSIRRLRALTKPIKCFKHVLFYQGCamide 和与果蝇、烟草天蛾的 CCAP 相比，在成熟肽的中部存在扩展现象（图 3：B）。CAPA 基因表达时进行了选择性剪接，对剪接产物 CAPA-a 和 CAPA-b 成熟肽的预测表明 CAPA-b 的成熟肽之一 CAPA-PK 与 CAPA-a 的成熟肽之一 CAPA-PK 相比也有所扩展（图 3：C）。

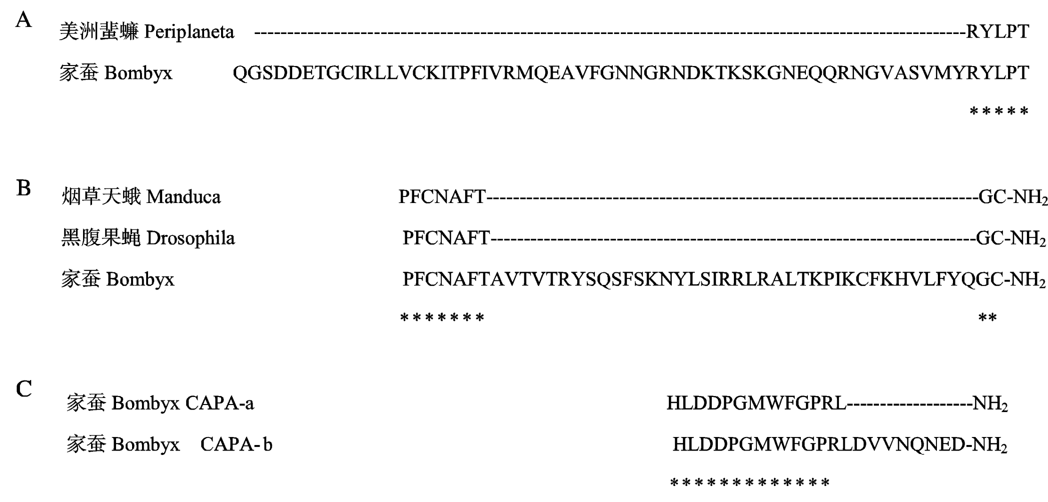


图 3 家蚕成熟神经肽结构的扩展性

Fig. 3 Structure extensibility of mature neuropeptides in *Bombyx mori*

A：Proctolin；B：CCAP；C：CAPA-PK. ＊代表序列相同；$-NH_2$ 代表 C 末端酰胺化。Asterisk represents sequence identity；$-NH_2$ represents the amidation of glycine residues in the figure.

3 讨论

分离纯化相应的神经肽是至今为止的主要方法。但基于取材和研究方法的难度，对神经肽的发现和研究进展较慢，且都限于对单个神经肽或基因的研究。随着家蚕基因组测序的完成，理论蛋白质数据库的建立，以及生物信息分析技术突飞猛进地发展，为神经肽基因快速筛查奠定了良好的基础。本研究利用已研究报道的昆虫神经肽与家蚕神经肽之间的同源性对家蚕基因数据库进行了全面而系统的神经肽基因筛查和成熟肽的预测，共获得 31 个神经肽基因家族，具体为 37 个神经肽基因亚家族，以及 44 个神经肽基因，并通过软件分析结合与已知昆虫或无脊椎动物神经肽的同源性，预测出 193 个成熟的神经肽。从上述的分析结果可见，家蚕神经肽前体通常由 90 ~ 250 个氨基酸残基组成，成熟神经肽的大小通常为 3 ~ 50 个氨基酸，与其他昆虫上的相关报道基本一致。对于本研究中多个预测的成熟神经肽序列长达 70 个氨基酸，甚至 100 多个氨基酸则需要进一步用实验进行验证。神经肽基因在高度保留家族的保守性特征序列的同时，在昆虫进化的过程中也进行了结构上的扩展。本项研究结果极大地扩充了家蚕神经肽的容量，为神经肽的功

能研究奠定了基础。

神经肽前体的加工是一个复杂的过程，包括多种翻译后的修饰方式。但预测得到的成熟肽的质荷比一方面可为质谱研究提供数据支持，尤其是使用个别神经系统进行直接的 MAL-DI-FTMS，可以评定预测的结构是否正确，同时也能发现一些非典型性的神经肽前体加工方式。另外本研究只能筛查到在昆虫或无脊椎动物中已经研究报道过的神经肽同源物，而对于尚未研究发现的家蚕中的神经肽或家族仍需进一步展开研究。

参 考 文 献（References）

Altschul SF, Madden TL, Schäffer AA, Zhang JH, Zhang Z, Miller W, Lipman DJ, 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research*, 25(17): 3 389 –3 402.

Bendtsen JD, Nielsen H, von Heijne G, Brunak S, 2004. Improved prediction of signal peptides: Signal P3.0. *J. Mol. Biol.*, 340: 783 –795.

Christie AE, Cashman CR, Brennan HR, Ma M, Sousa GL, Li L, Stemmler EA, Dickinson PS, 2008. Identification of putative crustacean neuropeptides using *in silico* analyses of publicly accessible expressed sequence tags. *Gen. Comp. Endocrinol.*, 156: 246 –264.

Coates D, Siviter R, Isaac RE, 2000. Exploring the *Caenorhabditis elegans* and *Drosophila melanogaster* genomes to understand neuropeptide and peptidase function. *Biochem. Soc. Trans.*, 28: 464 –469.

Dai L, Zitnan D, Adams ME, 2007. Strategic expression of ion transport peptide gene products in central and peripheral neurons of insects. *J. Comp. Neurol.*, 500: 353 –367.

Eipper BA, Mains RE, Herbert E, 1986. Peptides in the nervous system. *Trends Neurosci.*, 10: 463 –468.

Fernlund P, 1974. Structure of the red-pigment-concentrating hormone of the shrimp, *Pandalus borealis. Biochim. Biophys. Acta*, 371: 304 –311.

Gasteiger E, Gattiker A, Hoogland C, Ivanyi I, Appel RD, Bairoch A, 2003. ExPASy: the proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Research*, 31: 3 784 –3 788.

Hauser F, Cazzamali G, Williamson M, Park Y, Li B, Tanaka Y, Predel R, Neupert S, Schachtner J, Verleyen P, Grimmelikhuijzen CJP, 2008. A genome-wide inventory of neurohormone GPCRs in the red flour beetle *Tribolium castaneum. Front. Neuroendocrinol.*, 29: 142 –165.

Isaac RE, Taylor CA, Hamasaka Y, Nässel DR, Shirras AD, 2004. Proctolin in the post-genomic era: New insights and challenges. *Invert. Neurosci.*, 5: 51 –64.

Ishizaki H, Suzuki A, 1992. Brain secretory peptides of the silkmoth *Bombyx mori*: Prothoracicotropic hormone and bombyxin. *Progress in Brain Research*, 92: 1 –14.

Kastin AJ, 2006. Handbook of Biologically Active Peptides. Academic Press, New York.

Katoh K, Toh H, 2008. Improved accuracy of multiple ncRNA alignment by incorporating structural information into a MAFFT-based framework. *BMC Bioinformatics*, 9: 212.

Kramer SJ, Toschi A, Miller CA, Kataoka H, Quistad GB, Li JP, Carney RL, Schooley DA, 1991. Identification of an allatostatin from the tobacco hornworm *Manduca sexta. Proc. Natl. Acad. Sci. USA*, 88: 9 458 –9 462.

Monigatti F, Gasteiger E, Bairoch A, Jung E, 2002. The sulfinator: Predicting tyrosine sulfation sites in protein sequences. *Bioinformatics*, 18: 769 –770.

Pascual N, Castresana J, Valero ML, Andreu D, Bellés X, 2004. Orcokinins in insects and other invertebrates. *Insect Biochem. Mol. Biol.*, 34: 1 141 –1 146.

Sigrist CJA, Cerutti L, Hulo N, Gattiker A, Falquet L, Pagni M, Bairoch A, Bucher P, 2002. PROSITE: a documented database using patterns and profiles as motif descriptors. *Brief Bioinform.*, 3: 265 –274.

Southey BR, Amare A, Zimmerman TA, Rodriguez-Zas SL, Sweedler JV, 2006. NeuroPred: a tool to predict cleavage sites in neuropeptide precursors and provide the masses of the resulting peptides. *Nucleic Acids Research*, 34: 267 –272.

Stangier J, Hilbich C, Beyreuther K, Keller R, 1987. Unusual cardioactive peptide (CCAP) from pericardial organs of the shore crab *Carcinus maenas. Proc. Natl. Acad. Sci. USA*, 84: 575 –579.

Stangier J, Hilbich C, Burdzik S, Keller R, 1992. Orcokinin: a novel myotropic peptide from the nervous system of the crayfish, *Orconectes limosus. Peptides*, 13: 859 –864.

Starratt AN, Brown BE, 1975. Structure of the pentapeptide proctolin, a proposed neurotransmitter in insects. *Life Sci.*, 17: 1 253 –1 256.

Stemmler EA, Cashman CR, Messinger DI, Gardner NP, Dickinson PS, Christie AE, 2007. High-mass-resolution direct-tissue MALDI-FTMS reveals broad conservation of three neuropeptides (APSGFLGMRamide, GYRKPPFNGSIFamide and pQDLDHVFLRFamide) across members of seven decapod crustacean infraorders. *Peptides*, 28: 2 104 –2 115.

Veenstra JA, 1989. Isolation and structure of corazonin, a cardioactive peptide from the American cockroach. *FEBS Lett.*, 250: 231 –234.

Veenstra JA, 2000. Mono-and dibasic proteolytic cleavage sites in insect neuroendocrine peptide precursors. *Arch. Insect Biochem. Physiol.*, 43: 49 –63.

Williamson M, Lenz C, Winther AM, Nässel DR, Grimmelikhuijzen CJ, 2001. Molecular cloning, genomic organization, and expression of a C-type (*Manduca sexta*-type) allatostatin preprohormone from *Drosophila melanogaster. Biochem. Biophys. Res. Commun.*, 282: 124 –130.

Yoshida I, Moto K, Sakurai S, Iwami M, 1998. A novel member of the bombyxin gene family: Structure and expression of bombyxin G1 gene, an insulin-related peptide gene of the silkmoth *Bombyx mori. Dev. Genes Evol.*, 208: 407 –410.

（责任编辑：赵利辉）